

# WEEKLY REPORT

2014.04.21 – 2014.05.04

黄芯芯

## 近两周工作：

1. Palantir 项目调研，查找并研究一些多源数据集成的开源工具。我找到一个南加州大学开发的 Web-Karma 工具，可以集成从数据库、文本数据（csv、tsv 等）或者网络中得到的数据，并根据现有的 owl 文件手动对数据生成一个本体结构。我在本机上配置好了 Web-Karma 的环境并在周讨论中进行演示和讲解。另外，也找了一些爬去网络数据的开源工具库。
2. 配置浙一的插件开发环境。要使用 Java 版本的 vtk 进行开发。
3. 这两天写了体育新闻可视化的论文。
4. 下午与数媒的同学讲述项目背景，并给他们指示了几个可以回去讨论的点：
  - 现有的工作对球员在场上发生的时间只是简单的用不同的矩形进行编码，有没有更好的方式？
  - 一场比赛，球员与球员之间存在各种联系，如助攻、抢断和盖帽等，现有的工作中没有展示这些联系，设计一个展示球员联系的可视化

另外，今天下午有个新的 idea：之前一直想要通过某种计算方式来计算一场比赛中各个时间点的精彩程度，或者球员在各个时间点上的表现，然而通过公式来计算的结果误差肯定非常大，在今天下午的讨论上突然受到启发，可以通过协同的方式来判断比赛的精彩程度和球员的表现曲线。

当用户正在看比赛视频直播时，看到精彩的部分或者球员有很好的表现，用户可以在我发布的该场比赛的分差走势曲线上进行标注，即在该时间点上“点赞”或“踩”，或者对某个球员在某个时间点上“点赞”或“踩”，这样当多个用户在其上进行标注时，就可以累积得到一条时序的比赛精彩程度曲线，以及各个球员的时序的表现曲线。而没有看过直播的用户可以通过我们的可视化视图来定位到精彩的或者他感兴趣的时间段。

除了“点赞”或“踩”之外，用户也可以对各个时间点进行文本或者转折点标注等，后续的用户看到先前用户的标注可以发现比赛的转折点或精彩过程。

### 下周计划：

1. 继续 Palantir 项目调研。寻找并学习一些数据集成开源工具的源代码。
2. 指导数媒的同学们，定下具体方案，改进 NBA 项目。
3. 继续配置浙一插件开发环境。上周还没配好。。。